

EVALUATION OF THE SPACE SHUTTLE COCKPIT AVIONICS UPGRADE (CAU) DISPLAYS

Jeffrey W. McCandless, Ph.D. and Robert S. McCann, Ph.D.
NASA Ames Research Center
Moffett Field, CA

Kelsey W. Berumen, Stephen S. Gauvain, Victoria J. Palmer and William D. Stahl
United Space Alliance, LLC
Houston, TX

Andrew S. Hamilton
NASA Johnson Space Center
Houston, TX

In 1999, NASA Johnson Space Center (JSC) initiated a Cockpit Avionics Upgrade (CAU) project to increase shuttle mission safety by improving the crew's situation awareness, reducing their workload, and improving their performance. The primary focus of the project was a complete redesign of the current cockpit displays. To determine the effectiveness of the redesigns, a human-centered evaluation was conducted in the Shuttle Mission Simulator (SMS) at NASA JSC in 2003 and 2004. Measures of crew situation awareness, workload, and nominal and off-nominal task handling performance were made in a series of simulations of off-nominal flight situations during the dynamic flight phases of ascent and entry. The redesigned display formats yielded dramatic increases in situation awareness, reductions in workload, and improvements in performance.

1. INTRODUCTION

The Space Shuttle was developed in the 1970s using technology that was advanced for its time, including fly-by-wire components and multiple Cathode Ray Tube (CRT) screens in the cockpit. Although the electro-mechanical gauges and CRT screens soon became dated, no major upgrades were made to the cockpit for two decades. In 2000, Space Shuttle Atlantis made its first flight with the Multifunction Electronic Display System (MEDS) cockpit, which helped remedy the obsolescence and maintenance issues of the electro-mechanical gauges, dials, and CRT screens by replacing them with Liquid Crystal Display (LCD) equivalents. However, the MEDS upgrade (Figure 1) did not address serious human factors and usability deficiencies of the "legacy" cockpit displays. These deficiencies include:

- Much of the information is in monochrome (green on dark background) digital form arranged in closely spaced rows and columns. These characteristics make it difficult to localize and process key sources of information, such as off-nominal values.
- Information about one system or parameter is frequently scattered across two or more display formats, not all of which can be viewed at any one time. Consequently, extensive display navigation is needed to build a complete understanding of a system's current status and functional mode.

- On a single display, adjacent information might be referring to two independent systems, increasing the workload to decipher data.
- Key status and safety information is not available on any current display, forcing the crew to rely on communications with mission control to acquire the information.



Figure 1: MEDS cockpit

To address these deficiencies, the NASA Administration developed a usability-oriented modification called the CAU. The goal of the CAU project was to redesign and enhance the crew interface to improve situation awareness and performance, and reduce workload.

This paper presents the results of an evaluation to determine the efficiency of the CAU display redesigns. The evaluation consisted of scripted, formal evaluations comprising systematic simulations of various flight scenarios (including ascents, ascent aborts, and entries). The task of the crewmembers/astronauts was to perform nominal operations, detect, identify, and resolve simulated systems malfunctions, and perform mission impact assessments and other decisions relating to the malfunctions. Data were collected on situation awareness, workload, and performance to provide an indication of the impact of the proposed displays.

2. DESIGN PHILOSOPHY

In order to ensure a cockpit that was oriented around the needs of the flight crew, the designers for the displays were primarily from the operational community, which included astronauts, instructors, and flight controllers as well as human factors specialists.

2.1 Display Content

The foundation for the CAU project was a set of requirements intended to address the cockpit deficiencies. Implementation of the requirements guided the development of intuitive task-oriented displays, enhanced display applications, and future system growth to meet the objectives of the CAU (McCandless, McCann, and Hilty, 2003). The requirements are:

- Allow for multi-color graphics with data from multiple sources with logical information and command groupings on any display
- Expand avionics processing power for implementation of display applications/logic, plus reserve for future enhancements. Embedded computations were implemented to enhance the presentation of the displayed data
- Allow for tailoring of display information and command content to the current flight phase
- Allow for navigation to any display via existing display edge keys and keyboards

2.1.1 Color

CAU developers recognized that one of the biggest underutilizations of the glass LCD screens in the MEDS cockpit is the limited use of color. The proposed display formats use color in a systematic fashion to enable the crew to differentiate the varied classes of data and information, particularly during off-nominal conditions, defined as malfunctions (e.g., coolant leaks). Each color was specifically chosen based on display constraints and usability principles. For example, failure conditions were represented by four colors: red, yellow, orange and cyan. The critical colors of red

and yellow correspond to conventional meanings where red equals warning and yellow equals caution. The purpose of conventional coding for caution and warning colors is to draw attention rapidly. Orange represents a unique condition in which the primary and back-up computer systems produce different outputs (for example, if the primary computer system commands the engine thrust to a different level than the back-up computer system). Cyan represents cases in which data are unavailable for display because, for example, a sensor has failed.

2.2 Display Samples

2.2.1 MEDS

An example of the Horizontal Situation MEDS display is shown in Figure 2. It depicts an overhead view of the relationship of the orbiter and the runway during the final stages of re-entry.



Figure 2: MEDS Horizontal Situation Entry Display

The shuttle symbol in the center of the display is in a fixed position. The three dots in front of the shuttle symbol predict where the shuttle will be in 20, 40, and 60 seconds based on navigation state. The runway graphic (the circle with the line in it) represents the direction of the landing runway and the heading alignment cone. In this display example, the runway graphic shows the runway off to the left of the orbiter and the orbiter banking left (the 3 dots arc to the left) to approach it. Notably absent from this display is any indication of whether the shuttle has enough "energy" (i.e., speed and altitude) to make the runway.

2.2.2 CAU

The CAU design of the Horizontal Situation Display is shown in Figure 3. The new display format eliminates clutter, makes extensive use of color-coding, and adds some key

pieces of flight-related information to enhance situation awareness.

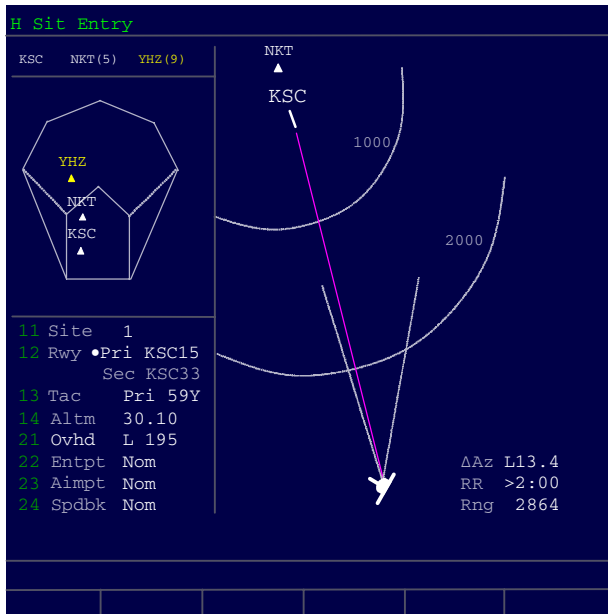


Figure 3: CAU Horizontal Situation Entry Display

The upper left corner provides an estimate of the orbiter's energy state with respect to potential landing sites ("footprint"). If a runway icon is in the middle section shaped like a house (e.g. KSC, NKT), the orbiter can reach that runway using normal flying techniques. If a runway icon is in the top region (e.g. YHZ), or the two triangular side regions, it is colored yellow and indicates that the orbiter does not have enough energy to make the runway without employing special flying techniques. If a runway is unachievable, it is colored red and depicted outside the footprint. Since the orbiter is simply a glider during the entry phase of flight, this information is extremely valuable to the on-board crewmembers.

The display also provides an indication of when the orbiter will next command a "roll reversal" to initiate one in a series of turns performed during entry to control energy. Two gray lines extend from the orbiter graphic indicating at what point the shuttle will command a bank from the left to the right (or vice versa).

3. TESTING METHOD

3.1 Experimental Conditions

The evaluation was conducted in the motion-based SMS at NASA JSC and utilized six crews of three astronauts each (Commander in the left seat, Pilot in the right seat, and Mission Specialist in the center aft seat). Each crew participated in both MEDS and CAU testing. The availability of the astronaut subjects and CAU hardware and software

dictated the testing schedule. During MEDS testing (October - December 2003), each crew completed three data collection sessions. Each session lasted approximately two hours and consisted of eight runs. The runs consisted of short (approximately 10 minute) time slices of a given flight phase. Each run contained several malfunctions and nominal tasks. At the end of each run, each crewmember independently filled out a questionnaire consisting of questions from the run just performed. Approximately 10 months later (August-September, 2004), the same crews completed three additional data collection sessions in the CAU cockpit. These sessions were identical to the MEDS sessions except that the simulator and procedures had been modified to support the CAU system. The purpose of the lengthy interlude between MEDS and CAU testing was to reduce the chances that the 18 astronaut participants would remember the exact details of the runs during the testing on the CAU displays. No additional formal training was provided to the test subjects during the interlude. A between subject design was rejected because only 18 astronauts were available for the study.

3.2 Types of data

3.2.1 Situation Awareness

The situation awareness data can be divided into objective and subjective categories:

1. Objective - Questions that had a definitive correct or incorrect answer. An example of an objective question is:

What was your trajectory energy state at the beginning of the run (circle one)?

High Low Nominal Don't Know

2. Subjective - Ratings based on crewmembers' opinions with no correct or incorrect answer. An example of a subjective question is:

Rate your situational awareness of the fuel cell problem as provided by the displays.

1 2 3 4 5 6 7 8 9 10

Several objective and subjective questions related to the scenarios just presented were answered by each crewmember following each run.

3.2.2 Workload

Workload measures for the evaluation are subjective. The two means utilized to measure workload were the Bedford Scale (Roscoe, 1984) and a modified version of the NASA Task Load Index (TLX) (Hart, 1988). Workload measures were collected from each crewmember at the completion of each run.

3.2.3 Crew Performance

Performance measurements were purely objective and centered around various aspects of crew behavior associated with malfunction detection, isolation, and recovery operations. These measures included the percentage of malfunctions that the crew failed to identify, the latency to recognize malfunctions that were identified, and errors in the fault management process following a correct identification. These errors included failure to initiate a remediation procedure, failure to complete a procedure, performing steps in a procedure incorrectly or out of order, and performing an incorrect procedure. Observers in the cockpit, as well as team members in the instructor station, recorded data on unidentified malfunctions and recognition times.

4. RESULTS

In the following tables, CDR corresponds with the Commander, PLT corresponds with the Pilot, and MS corresponds with the Mission Specialist. The numbers in the “Objective” rows refer to the percentage of questions answered correctly. The numbers in the “Subjective” rows refer to subjective ratings on a 1-10 scale. The numbers in the “Improvement” row were computed as (CAU-MEDS)/MEDS. Because of rounding, not all percentages appear exact.

4.1 Situation Awareness

The objective data for situation awareness of trajectory and orbiter systems indicate improvement from MEDS to CAU displays across all three crewmembers, as shown in Tables 1, 2 and 3. The subjective results also indicate substantial improvement.

For analysis of the situation awareness data, the data were broken into 3 distinct categories:

- A - Monitoring Trajectory
- B - Monitoring Data Processing System, Electrical Power System, Main Propulsion System and Orbital Maneuvering System
- C - Monitoring Auxiliary Power Units, Control, Environmental Control and Life Support System, Navigation and Reaction Control System

Statistical analysis (student's two-tail paired t-test for means) of the results was performed and proved that each of the results was met with at least 98% confidence for objective and subjective results. For each case (A, B and C), $df=17$ and $t\text{-critical} = 2.1$. For category A, $t\text{-stat}$ was 4.1 (for objective measures) and 8.4 (for subjective measures). For category B, $t\text{-stat}$ was 2.6 (objective) and 11.7 (subjective). For category C, $t\text{-stat}$ was 10.3 (objective) and 11.0 (subjective).

Table 1: Category A: Monitoring Trajectory

	Type	CDR	PLT	MS	Mean
Objective	MEDS	28%	21%	18%	22%
Objective	CAU	86%	73%	68%	76%
Improvement		203%	252%	286%	240%
Subjective	MEDS	3.2	3.2	3.2	3.2
Subjective	CAU	8.7	8.1	8.2	8.3
Improvement		169%	151%	158%	160%

Table 2: Category B: Monitoring Data Processing System, Electrical Power System, Main Propulsion System and Orbital Maneuvering System

	Type	CDR	PLT	MS	Mean
Objective	MEDS	46%	46%	32%	42%
Objective	CAU	77%	69%	74%	73%
Improvement		65%	49%	127%	75%
Subjective	MEDS	3.6	3.6	3.4	3.6
Subjective	CAU	8.6	8.1	8.1	8.3
Improvement		137%	123%	137%	132%

Table 3: Category C: Monitoring Auxiliary Power Units, Flight Control, Environmental Control and Life Support System, Navigation and Reaction Control System

	Type	CDR	PLT	MS	Mean
Objective	MEDS	23%	21%	17%	21%
Objective	CAU	69%	61%	77%	69%
Improvement		194%	188%	351%	235%
Subjective	MEDS	3.7	3.7	3.7	3.7
Subjective	CAU	8.4	7.9	7.9	8.1
Improvement		126%	118%	114%	119%

4.2 Workload

4.2.1 Bedford Scale

The results for the Bedford Scale show a significant drop in workload with the CAU displays and are consistent across all three crew positions, as shown in Table 4.

Statistical analysis (Student's two-tail paired t-test for means) of the crew results shows an improvement with more than 99% confidence for these subjective results ($t\text{-stat} = 10.4$, $df = 17$, $t\text{-critical} = 2.1$).

Table 4: Bedford Workload

Type	CDR	PLT	MS	Mean
MEDS	6.6	6.8	6.6	6.6
CAU	3.6	3.3	4.0	3.6
Improvement	46%	52%	39%	46%

4.2.2 NASA TLX

The MEDS and CAU results are quite similar for NASA TLX compared with the Bedford scale, as shown in Table 5. The advantage of TLX is that it breaks down workload into its subcomponents.

Table 5: NASA TLX Workload

	MD	PD	TD	E	F	Wgt Mean
MEDS	7.4	3.3	7.2	7.1	7.0	6.7
CAU	3.9	1.7	3.5	3.6	2.8	3.3
Improvement	47%	49%	52%	49%	60%	52%

(MD=Mental Demand, PD=Physical Demand, TD=Temporal Demand, E=Effort, F=Frustration)

4.3 Performance

4.3.1 Unidentified Malfunctions and Recognition Time

Crewmembers failed to identify malfunctions more frequently in the MEDS cockpit than the CAU cockpit, as shown in Table 6.

Additional analysis shows that 76% of the 153 malfunctions that were not identified with MEDS displays were identified with CAU displays. Conversely, only 29% of the 49 malfunctions that were not identified with CAU displays were identified with MEDS displays. Of those 29%, none was repeated, demonstrating that each of the cases was unique and each appeared to be an isolated incident.

Table 6: Unidentified Malfunctions

Type	Unidentified malfunctions
MEDS	30%
CAU	10%
Improvement	67%

4.3.2 Errors

Crewmembers are more likely to initiate and complete the correct procedure with CAU displays, and they are less likely to make errors, as shown in Table 7.

Table 7: Errors (percentage of total)

	Failed to initiate correct procedure	Procedures containing at least one error	Failed to complete correct procedure
MEDS	19%	22%	41%
CAU	10%	14%	25%
Improvement	49%	38%	39%

5. DISCUSSION

These results provide a comprehensive means of showing how the CAU displays positively affect the crew's situation awareness, workload and performance. The analysis is thorough in that it is based on literally thousands of distinct data points. In all cases, the results from the different sources and the different classes (objective and subjective measurements) are consistent with one another.

The CAU design process undoubtedly benefited from having carefully selected teams which developed each display. The design process included not only inputs from the end-user (i.e., astronauts), but also inputs from the software programmers and requirements analysts who provided a realistic indication of the feasibility of implementing the proposed designs. Others who provided useful input included astronaut instructors and flight controllers who possessed operational experience, and human factors specialists familiar with fundamental design principles. By ensuring that all relevant parties were involved in the display design, the final result was a markedly improved suite of displays.

The analysis shows that the CAU is an improvement over MEDS in the tested environment. The CAU displays provide an effective means of increasing the crew's safety. In spite of the apparent benefits of CAU, the project was cancelled in late 2004. Budget constraints limited the number of upgrades that will actually be implemented in the vehicle. Funding for the shuttle upgrades will instead be directed at other relevant areas. There are currently no plans to implement the proposed CAU displays. Nevertheless, the principles learned during this multi-year development and evaluation effort are still relevant as NASA presses forward to implement The Vision for Space Exploration (2004). The next-generation vehicle will likely contain a host of displays that depict vast quantities of systems information. The crews will benefit from displays employing the same guidelines used in CAU development. For this reason, the project can be considered successful.

6. ACKNOWLEDGMENTS

The Space Shuttle Cockpit Avionics Upgrade (CAU) project was based at NASA Johnson Space Center with support from other organizations such as NASA Ames Research Center.

Copyright © 2005 by United Space Alliance, LLC. Published by the Human Factors and Ergonomics Society with permission. These materials are sponsored by the National Aeronautics and Space Administration under Contract NAS9-20000. The U.S. Government retains a paid-up, nonexclusive, irrevocable worldwide license in such materials to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the U.S. Government. All other rights are reserved by the copyright owner.

7. REFERENCES

- Hart SG, 1988, Development of NASA-TLX Results of Empirical and Theoretical Research. In Hancock PA and Meshkati N (Eds) *Human Mental Workload*. North-Holland: Elsevier Science Publishers BV. 139-183.
- McCandless JW, McCann RS and Hilty BR (2003) Upgrades to the caution and warning system of the space shuttle. Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 16-20.
- Roscoe, AH (1984) Assessing pilot workload in flight. Advisory Group for Aerospace Research & Development Conference Proceedings No. 373: Flight Test Techniques. Neuilly-sur-Seine, France. NATO.
- The Vision for Space Exploration (2004). NASA. Washington, DC.